United States Military Academy
West Point, New York 10996

AN EVALUATION OF JOINT AND SERVICE-SPECIFIC ADVERTISING EFFICIENCY FOR MILITARY RECRUITMENT

OPERATIONS RESEARCH CENTER
TECHNICAL REPORT 01-06-1

Major Michael J. Kwinn, Jr, Ph.D.
Analyst, Operations Research Center

Lieutenant Colonel Mark J. Davis
Director, Operations Research Center

Colonel Michael L. McGinnis, Ph.D.
Professor and Head, Department of Systems Engineering

15 December 2000

The Operations Research Center is supported by the Assistant Secretary of the Army (Financial Management)

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*
15 December 2000

2. REPORT TYPE
Tech

3. DATES COVERED *(From - To)*
7/1/00 to 6/30/01

4. TITLE AND SUBTITLE
An Evaluation of Joint and Service-Specific Advertising Efficiency for Military Recruitment

5a. CONTRACT NUMBER
n/a

5b. GRANT NUMBER
n/a

5c. PROGRAM ELEMENT NUMBER
n/a

6. AUTHOR(S)
Major Michael J. Kwinn, Jr., Ph.D.
Lieutenant Colonel Mark J. Davis

5d. PROJECT NUMBER
n/a

5e. TASK NUMBER
n/a

5f. WORK UNIT NUMBER
n/a

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Operations Research Center of Excellence
Department of Systems Engineerig
United States Military Academy
West Point, New York 10996

8. PERFORMING ORGANIZATION REPORT NUMBER
01-06-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
United States Army Recruiting Command

10. SPONSOR/MONITOR'S ACRONYM(S)
USAREC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
n/a

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Statement A - Approved for Public Release; Distribution is Unlimited

13. SUPPLEMENTARY NOTES
n/a

14. ABSTRACT
This report examines whether advertising money is more efficiently allocated to Joint advertising or to Service-specific advertising (Army, Navy, Air Force, Marines). This is done using data gathered in 1984 under the Department of Defense sponsored "Advertising Mix Test" wherein a designed experiment varied the levels of joint and service-specific advertising across the US and observed the number of recruits obtained. Previous studies have not considered the efficiency with which different entities conduct recruiting activities, and it is possible that a good program can be inefficiently run, or an inferior program can be efficiently run, thus leading to incorrect conclusions if efficiency is ignored. Here we show that in the test data design, the "joint advertising" cells had 5-15 times as many efficient recruiting entities as had the "service specific advertising" cells, and that ignoring this efficiency difference leads to the conclusion that joint advertising is more efficient that service specific advertising. After removing managerial inefficiencies in each program, however, we arrive at exactly the opposite conclusion, namely that when efficiently managed service

15. SUBJECT TERMS
Recruiting Efficiency

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
n/a	n/a	n/a	SAR	30	LTC Mike Kwinn, Jr.

19b. TELEPHONE NUMBER *(Include area code)*
845-938-5591

AN EVALUATION OF JOINT AND SERVICE-SPECIFIC ADVERTISING EFFICIENCY FOR MILITARY RECRUITMENT

Lead Analyst
Major Michael J. Kwinn, Jr., Ph.D.
Analyst, Operations Research Center

A TECHNICAL REPORT
OF THE
OPERATIONS RESEARCH CENTER
UNITED STATES MILITARY ACADEMY

Directed by
Lieutenant Colonel Mark J. Davis
Director, Operations Research Center

Approved by
Colonel Michael L. Mc Ginnis, Ph.D.
Professor and Head, Department of Systems Engineering

15 December 2000

The Operations Research Center is supported by the Assistant Secretary of the Army (Financial Management)

The sponsor for this project is Plans and Analysis Division, EPMD, PERSCOM

About the Author

MAJOR MICHAEL J. KWINN, JR. is an Analyst in the Operations Research Center at the United States Military Academy (USMA), West Point, New York. He received a BS from USMA in 1984, an MS in Systems Engineering from the University of Arizona and a Ph.D. in Management Science and Information Systems from the University of Texas at Austin. His research interests are personnel management systems and force structure analysis through the use of optimization and simulation models.

Acknowledgments

I would like to acknowledge the Office of the Assistant Secretary of the Army for Financial Management (ASA-FM) for funding this project. I would also like to acknowledge some individuals who helped in the development of this approach and analysis:

Dr. David A. Thomas, a member of my committee and the individual who introduced me to this area of research.

Dr. Patrcik L. Brockett and Dr. William W. Cooper, my co-supervisors of my committee for this research.

Colonel Gregory Parlier, Director of Program Analysis and Evaluation, US Army Recruiting Command, for his interest in this line of research and for the promotion of this project in his organization.

Table of Contents

Executive Summary

This report examines whether advertising money is more efficiently allocated to Joint advertising or to Service-specific advertising (Army, Navy, Air Force, Marines). This is done using data gathered in 1984 under the Department of Defense sponsored "Advertising Mix Test" wherein a designed experiment varied the levels of joint and service-specific advertising across the US and observed the number of recruits obtained. Previous studies have not considered the efficiency with which different entities conduct recruiting activities, and it is possible that a good program can be inefficiently run, or an inferior program can be efficiently run, thus leading to incorrect conclusions if efficiency is ignored. Here we show that in the test data design, the "joint advertising" cells had 5-15 times as many efficient recruiting entities as had the "service specific advertising" cells, and that ignoring this efficiency difference leads to the conclusion that joint advertising is more efficient that service specific advertising. After removing managerial inefficiencies in each program, however, we arrive at exactly the opposite conclusion, namely that when efficiently managed service specific advertising is more efficient that is efficiently managed joint advertising.

1. Introduction

This technical report examines whether advertising activities directed toward military recruitment should be combined into a single "joint" effort or whether such advertising is better conducted, as is currently done, by the specific services (Army, Navy, Air Force, Marines). That this topic is important is evidenced by the numerous studies on this subject. In fact, the data we use in this analysis are drawn from a large, multi-million dollar study conducted by the Wharton Center for Applied Research (WCAR) in the 1980s with Department of Defense (DoD) support. Unfortunately, due to deficiencies in their analysis, the WCAR study did not resolve the issue (c.f., Charnes, Cooper, Golany and Brockett 1986). It led, instead, to other studies that also did not adequately address certain pertinent aspects of the issue.

Interest in this topic has again surfaced, at least in part, because of serious shortfalls in the number of persons being recruited for military service. Indeed, this topic has recently been the subject of presentations and briefings at policy levels of DoD. The latest discussion of the Joint advertising issue began again in earnest on August 4, 1999 when Bozell/Eskew Advertising presented an official briefing of their findings from their review of recruitment advertising in the DoD. In that briefing, one of their many recommendations was to "Increase [the Joint Recruiting Advertising Program (JRAP)] budget exponentially." They suggest returning to "historical" budget levels, much like the levels that the WCAR conducted. Advertising Mix Test sought to analyze. They suggested that with this increase in budget the Joint Recruiting Advertising Program (JRAP) could be charged with "corporate branding". However any actually empirical evidence to support this recommendation is absent. As the JRAP budget was reduced to

essentially zero in the late 1980's, the Advertising Mix data set we use from the WCAR study is the only means of validating or repudiating this assertion. This is done in this report.

One significant consideration necessary to attend to when attempting to rigorously deal with this problem is how to unravel the effect of differential efficiency with which different recruiting efforts are being conducted. As noted in Thomas (1990), variables such as the allocation and use of recruiting personnel as well as background environmental factors, such as varying rates of unemployment in different recruiting districts, need to be taken into account (along with advertising) when evaluating recruiting activities. That is, a complex collection of other activities, both discretionary and non-discretionary, can affect the numbers of personnel that are recruited. In such situations, one may easily imagine cases in which an intrinsically superior strategy for advertising in one district is inefficiently run and consequently is associated with performances that are inferior to an intrinsically less effective strategic advertising program that happened to be coupled with more efficient recruiting efforts in another recruitment district. Accordingly, without properly addressing the effects of differential efficiency in recruiting, one may be (and others have been) led seriously astray, and led to erroneous conclusions about the superiority or inferiority of joint versus service specific advertising programs – results which, we show, are reversed once the differential efficiency of administration is simultaneously considered and factored out. Without taking into account the efficiency or inefficiency of recruiting behavior, one confronts a confounded hypothesis test to analyze the merits of joint versus service specific advertising. How such efficiencies and inefficiencies are to be taken into account in

evaluating the effects of advertising strategies forms a major focus for this report, and is missing from the studies we examined on the topic (with the exception of Thomas (1990)). This report is the first to rectify this issue and separate effectiveness of joint versus service specific advertising as a part of a recruitment strategy from the efficiency of administration of the joint versus service specific advertising program, and accordingly this report is able to come to conclusive results where other papers have not, and explain why other studies have come to opposite conclusions by failing to factor out important factors.

To identify efficiency possibilities and shortcomings, we follow Thomas (1990) and utilize concepts and methods that are drawn from Data Envelopment Analysis (DEA). This use of DEA provides an alternative to the use of statistical regressions in analyzing joint versus service specific advertising, and is able to distinguish between efficient and inefficient performances in recruiting activities that need to be considered when evaluating the effects of military recruiting.

2. Prior Advertising Studies Related to Recruiting

Early in the history of the All Volunteer Force (AVF), government agencies conducted analyses of the effectiveness of the recruiting efforts (U.S. Government Reports 1976,1976). These reports were essentially developed at the macro level and addressed broad overall policy issues. They did not provide specific policy recommendations based on empirical work, as there was very little data on recruiting available under the AVF at that time. Since 1980 most of the important reports on military recruiting were conducted either by academia or by government-sponsored organizations such as the RAND Corporation. The majority of these reports studied

military recruiting through classical statistical econometric analysis techniques (e.g. used ordinary least squares regressions). Goldberg, for example, used this approach to study the impact of advertising on Navy recruiting (Goldberg, 1982). He also, (with Greenston, 1986) used econometric approaches to analyze Army enlistments. Dertouzos and Polich (Dertouzos and Polich, 1985) used an econometric approach to analyze of the effects of advertising on Army recruiting. In this report they also gave a detailed account of how advertising affects recruiting and extended earlier efforts by modeling the time-lag effects of advertising.

Daula and Smith (1986) also used an econometric approach. This study used a classical supply-and-demand approach as taken from economics to model enlistments. In this development of a demand model, they included recruiting goals. In their supply analysis, Daula and Smith used advertising "impressions" versus advertising expenditures. The measure of impressions they used was the total number of times an ad is seen by the target population. This is a commonly used measure in marketing/advertising analyses. Dertouzos (1985) expanded on the recruiter goals concept by defining in detail the pressures on recruiters both when they fail to meet, and when they exceed, established goals.

Dale and Gilroy (1986) attempted to model the effects of the state of the economy on enlistments. However, their use of classical statistical modeling approaches contains certain difficulties. For instance, they modeled the unemployment rate as a "discretionary variable". That is, unemployment was treated as a variable that can be varied at the discretion of management. This is not correct. It is better viewed as being an exogenously fixed variable to which management must adjust. Unemployment can

have a significant impact on recruiting and hence should be taken into account in manners that we will subsequently explore.

The RAND Corporation conducted a number of very important studies on recruiting throughout the 1980's and early 1990's, with many of the most significant studies during this time being led by James Dertouzos (1985, 1989). See also Dertouzos and Polich (1985, 1989). In the Army advertising study mentioned above, he used an ordinary least squares regression econometric approach to analyze data obtained in a three-year period (from 1981-83). He tried to distinguish between the effects of different broadcast and print media advertisements and also tried to determine the varied effects of national versus local advertising. Finally, he introduced a new level of sophistication by introducing a geometric decay function to model lagged effects of advertising.

Two other works have attempted to summarize the findings of much of the recruiting research conducted in the 1980s. Warner (1990) conducted his own analysis of the effects of recruiters, advertising and incentive programs, such as the Army College Fund. Using standard regressions techniques (he used two models: one included a time trend and one did not), he concluded that advertising and the number of recruiters has large effects on Army recruiting, but a lesser effect on recruiting for the other services. Warner also found that unemployment and relative pay differentials between the civilian and military work forces are the most significant factors in recruiting. Sohn (1996) used a random effects meta analysis technique to consolidate and summarize the results of many studies conducted in the 1980s. Based on the conclusions reached in four separate econometric studies, Sohn concluded from his meta analysis that increasing either the

recruiting force or the advertising budget in the Army would increase high quality recruits.

Evidently, a study of advertising effects on military recruitment can take many different forms and move in many different directions. Our focus is on Joint versus Service-Specific approaches to this kind of advertising. Hence this is an area in which prior work will be explained in more detail.

3. The Advertising Mix Test

The origin of the Advertising Mix Test can be traced to a Congressional Budget Office study released in 1981 that "recommended significant increases in the Joint Recruiting Advertising Program (JRAP) with concurrent reductions in Service-specific advertising and a net overall saving." (Korb memo in Carroll, 1987) The individual armed services were concerned about this proposal. Due to the complete lack of empirical support and the resistance of the services toward the recommended changes, Lawrence J. Korb, the Assistant Secretary of Defense (Manpower, Reserve Affairs and Logistics) commissioned the WCAR to conduct a study of immense proportions. The Department of Defense directed Advertising Mix Test (AMT) conducted in fiscal year 1984 involved every recruiting organization in the Continental United States (CONUS) for all services. The information gathering effort itself was monumental. This test remains the only publicly known nationwide market research effort of the type that essentially compared and controlled "category" versus "brand" advertising effectiveness in a "controlled design" experiment.

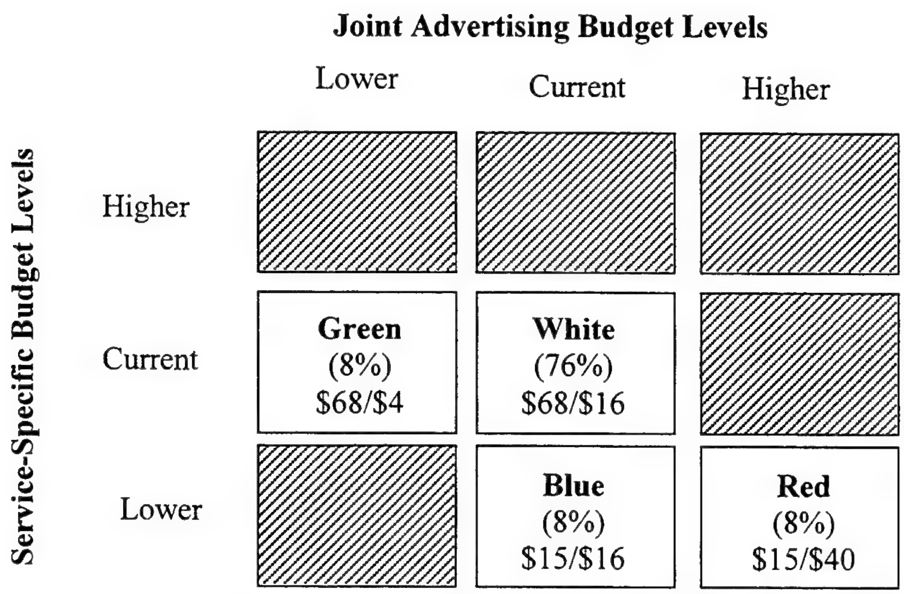
The test originally used an experimental design approach that was designed to vary the size of the joint and service-specific budgets across nine test cells. The budgets

in the test cells would be high, current or low for both joint advertising and service-specific advertising based on the FY1982 budget levels. After consulting with DoD officials, WCAR eliminated the three cells corresponding to an increase in the service-specific budgets and the one cell corresponding to a higher joint/current service-specific budget since these cells would have required a much higher overall budget and this could not occur without Congressional approval. The services themselves forced the elimination of the cell with lower spending levels for both joint and service-specific because they felt that effective recruiting sufficient to make their mission would be impossible in those areas.

The final cells, the percent of the 17-21 year old male population, and the budget levels are depicted at Figure 1, below. The Green Cell represents the (then) current (1982) level of Service Specific advertising funding ($68 million) and a lower level of Joint advertising funding ($4 million). The Blue Cell includes a lower level of Service Specific advertising funding ($15 million) and the current level of Joint advertising funding ($16 million). The Red Cell has a lower level of Service-Specific advertising funding (15 million) and a higher level of Joint advertising funding ($40 million). The White Cell (the "control" cell) includes the current levels of both Service-Specific and Joint advertising funding ($68 million and $16 million, respectively).

After designing the test, the WCAR contracted the RAND Corporation to assign geographic portions of the population into each of the cells so as to balance them. The criteria they used to balance the cells were "size of population, enlistment rates, unemployment and enlistment propensity" and finally geographical dispersion was included as an additional consideration. They placed in each cell, a number of

geographical regions known as Areas of Dominant Influence (ADIs) for the advertising media under study. These ADIs were developed by the Arbitron Ratings Company many years ago to help target advertising within different markets (the WCAR study used the partitions developed in 1981). ADIs are geographical areas that are predominantly covered by a specified local television-viewing pattern. The 212 ADIs within the United States completely partition the country into county groups. Figure 1, below, reflects these cell contents



Note: Cell White (the control cell) was eventually reduced to only a subset of the original ADIs. This subset used for analysis became Cell Yellow

Figure 1: Actual Cell Design

The percentages in each cell represent the percent of the 17-21 year old male population represented in each cell. The White Cell (the control) initially contained approximately 76% of the population. The other cells, Green, Blue and Red, were given approximately 8% of the population.

PEP Systems Incorporated received data from each of the services, which it consolidated and provided to WCAR. They compiled the information presented by the services into ADIs and then included the advertising expenditures. This created an

extremely large database. Unfortunately, the data were not all complete or logical (e.g. the Marine Corps reported "negative" recruiting goals in some periods), and almost from the beginning, there were problems with the test. The budgets were cut for many of the ADIs in the control cell as both the Joint Recruiting Advertising Program and the Navy reduced their budgets. WCAR decided to reduce the White cell to a subset of ADIs whose budgets had not been reduced. This subset of the control cell, which became the Yellow cell, contained only 16% of the population. Unfortunately, this cell (the original White cell) lost most of its "balancing" qualities that the original design of experiments attempted.

Upon obtaining the database from PEP Systems, Inc., WCAR began its analyses, using a simple log-linear regression model for their analysis with the dependent variable being the number of High Quality Male High School Graduate contracts. Key regressor variables were the number of recruiters, percent unemployment, percent urbanization, and advertising levels. Additionally, WCAR added a dummy variable with a value of 1 if the ADI was in a particular cell and 0 otherwise.

3.1. Critique of The WCAR Analysis of the Ad Mix Test

The test design, the implementation, the cell balancing and the test analysis have all been examined and critiqued with comprehensive critiques provided by the Center for Cybernetic Studies (CCS) at the University of Texas (Charnes, et al, 1986) and the RAND Corporation (Dertouzos, 1989). Both of these organizations conducted their own studies using the original or subsets of the original data sets partly because each experienced difficulty in replicating the WCAR results. The reader is referred to these studies for a more complete critique of the WCAR study.

WCAR declared the Blue cell as the "winner" of the test (the cell with the lower service and current joint levels.). Unfortunately, this conclusion was not supported by the analysis since most of the regression coefficients used to reach this conclusion were not significantly different from zero, even at the 90% confidence level, and the Red cell actually had more recruit contracts per capita. Additionally, WCAR did not even discuss the impact of recruiters, though this variable had a higher coefficient than did advertising in their modeling efforts.

In spite of this lack of statistical support for their conclusion, in their final analysis, WCAR makes some very dramatic policy recommendations. They conclude, "The size of the Joint advertising budget should be increased as Service-specific budgets are scaled back…[and] the Department of Defense can reduce its total advertising spending without adversely effecting recruiting performance."

3.2. Other Critiques of the WCAR Results

Naturally, the services immediately scrambled to verify or refute the report and its findings. The Army, with the biggest budget and the most to lose, was the first to critique the findings. The United States Army Recruiting Command (USAREC) began work with the Center for Cybernetic Studies (CCS) at the University of Texas at Austin. As CCS could not obtain the data set from WCAR, they compiled a similar data set using the same time frame through the efforts of the Defense Manpower Data Center (DMDC) and the USAREC. Their research concentrated only on Army effects of this test. They used Joint advertising expense information, but did not analyze the effects on other services. They used all 211 ADIs (one ADI was not used as it had incomplete and useless data) and therefore obtained much more complete and robust results than the WCAR study.

CCS concluded that the Department of Defense should not use the WCAR study policy recommendations because the methodology and the basis for conclusions in the WCAR study were flawed (c.f., Charnes, et al 1986).

The Department of Defense, faced with great criticism of the WCAR study from all services, hired the RAND Corporation to determine which study was correct, the WCAR study or the CCS study. In their work, RAND attempted to replicate the WCAR study to determine some of the problems with it. RAND discussed many of the same faults that the CCS study found in their analysis of the WCAR work, but did not address the CCS study at all. They conducted a separate study of the dataset using standard econometric models. RAND's conclusions were to not use the WCAR study and that their "results do not provide unequivocal conclusions about the relative efficacy of joint versus service [-specific] programs" see Dertouzos, 1989)[1].

One of the deficiencies of all of the previous analyses of joint versus service specific advertising effectiveness is that the ultimate outcome of the recruiting endeavor depends highly upon the efficiency with which all of the elements of the endeavor (recruiters, advertising, base target population size, unemployment rate, etc.) are all put together. No previous study has attempted to address this critical issue. Accordingly, we now undertake the efficiency study needed showing how such efficiency can be deduced, and then applied.

4. DEA and Assessing the Efficiency of DMUs

We begin with a general description of DEA and its applications. Since eliminating or controlling for the confounding effects of differences in efficiency in the

[1] For a more complete discussion of the deficiencies of the RAND study, see Kwinn 2000.

management of recruiting in the Advertising Mix study cells is critical to being able to make more definitive judgments concerning the superiority or inferiority of joint advertising, this will put us in position to apply these concepts and methods to the Advertising Mix Test data.

DEA originated with Charnes, Cooper and Rhodes (1978) generalizing the notion of efficiency used in engineering and economics for single input-single output production. It uses a mathematical programming method to locate an empirical "efficiency frontier" that can then be used to evaluate the relative efficiency of each entity considered for comparison. DEA refers to the entities being evaluated as “Decision Making Units” or DMUs for short. The term DMU is generic in character. It can refer to easily identified entities such as business firms or district recruiting offices. It can also refer to regions, areas, or countries, as well as subdivisions of entities (such as surgical units in hospitals), which utilize the same kinds of inputs to produce the same kinds of outputs according to the framework of the problem under analysis. Since its inception, DEA research has extended its domain of application considerably. Books on the subject include those by Ganley and Cubbin (1992), Norman and Stoker (1991) and Cooper, Seiford, and Tone (1999). It is now perhaps the most widely accepted method for analysis of efficiency of production, and the mathematical details of the methodology can be found in the books above and will not be reproduced here.

Briefly, DEA identifies those DMUs that utilize a minimum number of inputs to produce a given level of outputs. The piecewise linear function constructed by connecting the output values corresponding to the input values of the efficient DMUs constitutes the efficient production frontier function. A DMU whose exhibited

production is not on this derived efficiency frontier is termed inefficient. As a byproduct of the mathematical programming approach to efficiency analysis, DEA explicitly identifies the sources and amounts of inefficiency for the inefficient units and also gives a summary measure of relative efficiency for each unit, and allows the construction of an empirical production function that is determined by the efficient DMUs. This determination of the production function via an extremal process at the individual DMU level of analysis is in contrast to regression-based methodologies that determine average or typical production which might be expected by the collection of DMUs but does not focus on most efficient production. The set of efficient DMUs identified by DEA determine a "best practice (as opposed to average or typical) production frontier that can then be used for econometric analysis or statistical tests with technically inefficient production removed. In addition, each inefficient unit can be "projected" to the efficient frontier to determine what the output should be for the inefficient unit using their own inputs and calculating what would be their efficient level of output.

The ability to consider efficient production makes it possible to distinguish between the efficiencies of a program or strategy (e.g., an advertising strategy such as joint versus service specific) as distinct from the way in which it is managed. Thus, a comparison of the program or strategy possibilities can be conducted after these "managerial inefficiencies" were removed. This makes it possible to evaluate the potential of a program or strategy as distinct from the inefficiencies that might be present in the way in which it is operated.

We will subsequently use this way of separating efficient from inefficient behavior and apply it to the different recruitment districts, or rather Areas of Dominant

Influence (ADIs) involved in our study of the way advertising is used (in combination with other resources) in military recruiting activities. ADI is a concept developed by the Arbitron Corporation to measure the effects of advertising over contiguous county regions where the local television stations are dominant. There were 212 of these areas at the time of the Advertising Mix Test.

5. Program Evaluation from a Statistical Perspective

Due to the nonparametric nature of DEA, it is natural to use nonparametric statistical tests to examine differences between Joint and Service Specific advertising efficiency (as measured by DEA). We start by delineating the inputs used in our DEA analysis. These are given in Table 1. We do the same analysis for each service, and shall use Army as the prototypical service for explication purposes. As can be seen, only Number of Recruiters, Quota, and Army Advertising are listed as being discretionary variables in our evaluation of Army performances (i.e., variables under the control of the Army for the purposes of managerial adjustment to affect efficiency. In particular, the Navy, Air Force, Marines and Joint advertising expenditures are all regarded as non-discretionary because, from the perspective of Army recruiting, the Army does not possess the authority to vary these expenditures.

Input	Categorization
17-21 y/o male Population	Non-discretionary
Number Unemployed	Non-discretionary
Total Income	Non-discretionary
Number of Recruiters	Discretionary
Quota	Discretionary
Joint Advertising	Non-discretionary
Army Advertising	Discretionary
Navy Advertising	Non-discretionary
Air Force Advertising	Non-discretionary
Marines Advertising	Non-discretionary

Table 1: Categorization of Input Variables for Army Analysis. A discretionary variable is one that is under the control of the DMU and can be adjusted by managerial actions to affect efficiency

The above table refers to the Army. Similar tables for the Navy, Air Force, and Marines would reflect this same kind of categorization for the non-discretionary and discretionary variables. Joint advertising, however, is accorded a different treatment. As noted earlier, the Bozell/Eskew report recommends that all advertising for recruitment be assigned to a new separately organized agency. Therefore, to test this part of their proposal, we treat all advertising inputs as discretionary in the case of Joint advertising. This shall become important later as projection of inefficient units to the efficiency frontier can occur only through adjustment of the discretionary variables in the analysis. As can be seen in Figure 1, spending on advertising in the Green and White Cells is predominantly Service-specific advertising. Conversely, advertising spending in the ADIs in the Blue and Red cells is predominantly Joint. Hence, for our analysis, the ADIs in the Green and White cells will represent the Service-specific advertising program and the ADIs in the Blue and Red cells will represent the Joint advertising program.

In order to evaluate the hypothesis that the two programs are equally efficient, we follow Brockett and Golany (1996) and use the Mann-Whitney rank order statistic in as follows:

1. Rank the DMUs by their relative efficiency rating as obtained from the combined efficiency analysis of the DMUs. All ties in efficiency ratings are assigned the mid-rank value.

2. Compute the sum of all rankings of one of the programs. This value is labeled R.

3. Use R to compute the Mann-Whitney rank statistic:

$$U = n_1 * n_2 + \frac{n_1 * (n_1 + 1)}{2} - R$$

where n_1 and n_2 are the number of observations in programs 1 and 2, respectively and $n = n_1 + n_2$.

4. For sufficiently large values of n_1 and n_2, compute:

$$Z_{test} = \frac{U - \frac{n_1 * n_2}{2}}{\sqrt{\frac{n_1 * n_2 * (n_1 + n_2 + 1)}{12}}}$$

which is approximately normally distributed.

As Z_{test} is approximately normal for $n_1, n_2 \geq 10$, we accept the null hypothesis that the programs are equally efficient if $-Z_{\alpha/2} \leq Z_{test} \leq Z_{\alpha/2}$. Alternatively, we can utilize a one-sided test by assuming that the hypothesis to be tested is that Program 1 is more efficient. The test is then formulated as $Z_{test} \geq Z_{\alpha/2}$, and we can reject the null hypothesis and conclude that Program 1 is more efficient than Program 2 when Z_{test} is significantly greater than $Z_{\alpha/2}$. Similarly, if $Z_{test} \leq -Z_{\alpha/2}$, we reject the null hypothesis and conclude that Program 2 is more efficient than Program 1.

5.1. An Analysis of Program Differences Without Removing Managerial Inefficiencies

To effect this analysis, the above described DEA was run and the DMUs are rank ordered according to their efficiency, and the sum of the values of the Service-specific program is used to develop a Mann-Whitney value, which is converted into a standard normal statistic and compared to the normal distribution for hypothesis testing. The hypothesis test can be stated as follows:

H_0: There is no difference between the two programs, versus

H_1: Service specific advertising is more efficient than Joint advertising, or Joint advertising is more efficient than Service-Specific advertising.

Notice that this is a two-sided hypothesis test. For Z-values greater than 1.96, we reject the null hypothesis and conclude that Service-specific advertising is more efficient than Joint advertising. For Z-values less than –1.96, we reject the null hypothesis and conclude that Joint advertising is more efficient than Service-specific advertising.

The results of our analyses are presented in Table 2. The number of observations in the Service-specific program is n_1 and the number in the Joint program is n_2. The sum of the ranks for the Service-specific observations is R. This value is used in the calculation of U, the Mann-Whitney rank statistic, and the standard normal statistic, Z, in order to perform our tests of these H_o and H_1 hypotheses.

	n_1	n_2	R	U	Z
Army	2148	384	2,811,258	321,600.5	**-6.88**
Navy	2148	384	2,819,765	313,093	**-7.53**
Air Force	2148	384	2,779,281	353,577	**-4.46**
Marine	2148	384	2,798,462	334,396.5	**-5.91**
All-Services	2148	384	2,810,194	322,664	**-6.80**

Table 2: Results from Programmatic Analysis without Removing Managerial Inefficiency

As we can see in the table, if we do not remove the managerial inefficiency from the recruiting operation, we reject the null hypothesis of equal efficiency with a high level of statistical significance, and conclude that Joint advertising is more efficient than Service-Specific advertising for each service, including the analysis of All-Services. One way of interpreting these results is as follows: if we assume that recruiting operations cannot be managed efficiently (or if we simply ignore any inefficiencies in recruiting), then Joint Advertising is more efficient. More precisely, this conclusion is warranted if the different degrees of inefficiencies in each ADI are not altered.

An alternative explanation of the above results is that while there are indeed differences between the Joint and Service Specific cells which result in recruitment performance differences, this differences is due to something other than advertising program differences, and something which was not controlled in the original experimental design. If such a confronting variable is found, then one must control for this variable and then reanalyze to see if the results still stand after control is in place. We do this now.

The most obvious candidate for alternatively explaining the above results is that there are recruiting administration efficiency differences between the Joint and Service

Specific cells. To examine this possibility, a calculation was done of the percentage of efficient DMUs in each program (Joint and Service Specific). The results are given below for each of the service specific DEA analyses previously described.

Services	Joint	Service Specific	Ratio: Joint/Service Specific
Army	0.3426	0.0483	7.09
Navy	0.4066	0.0815	4.99
Air Force	0.4941	0.0312	5.84
Marines	0.4436	0.0745	5.95
All Service	0.4490	0.0649	6.92

Table 3. Ratio of Efficient to Inefficient Performers: Joint and Service-Specific

As can be seen, the Joint cells were relatively "loaded" with managerial efficient units, having between about 5 to 15 times as many efficient units per analysis as did the Service Specific cells. Given the dramatic differences in managerial efficiency between the two programs, it is tempting to assert that the apparent dominance of the Joint Advertising is due to managerial efficiency and not due to programmatic efficiency. Any potential superiority of the advertising program is overwhelmed by the huge uncontrolled differences in managerial efficiency. Before being ale to make any definitive conclusions, one must account for the efficiency differences between programs. We do this now.

5.2. *An Analysis Using Only Efficient Performing Units*

We now turn to our treatment of advertising "program efficiency" as distinct from the ways in which the Service-Specific and Joint advertising programs are managed. Because the number of observations is sufficiently large, we can start here by simply confining our attention to the efficient performances that were actually observed in each program, and comparing them. Proceeding in this manner, for instance, we find that in

the Army analysis only $n_1 = 99$ Service-Specific DMUs (=ADIs) were fully efficient while $n_2 = 98$ DMUs were found to be efficient for the Joint advertising DMUs. Indeed, in all cases, the cells reporting Joint advertising exhibited a higher proportion of efficient performances, as documented in Table 3. Finally, for "All-Services" we have the efficiency proportion for Service-specific advertising represented by $131/2148 = 0.06$, while Joint advertising yields $119/386 = 0.35$.

Table 4 reports the results from applying the Mann-Whitney statistic to these data involving only the purely efficiently administered programs. All services, except the "All-Services" category yield positive values that are statistically significant. This shows Service-Specific advertising, when efficiently managed, is statistically superior to efficiently managed Joint.

	n_1	n_2	R	U	Z
Army	99	98	7,381.5	7,270.5	**6.05**
Navy	162	111	16,587	14,598	**8.75**
Air Force	176	127	18,448	19,480	**11.04**
Marine	149	118	14,243	14,514	**9.13**
All-Services	131	119	15,707	8,528	**1.28**

Table 4: Results from Programmatic Analysis using only Efficient Performers

We conclude that our analysis strongly supports the hypothesis that Service-Specific advertising is more efficient than Joint Advertising <u>for each service</u>. Hence, we conclude with high probability that the results favoring Joint advertising in Table 2 are due at least in part, to the much larger proportion of efficient ADIs in the cells for Joint advertising. In any event, the confounding of efficiency and advertising call into question any analysis that attributes superiority only to the latter.

As a second check on the above results we will use DEA to analyze the relative efficiency of Service-specific versus Joint advertising controlling for efficiency differences as found previously. In our analysis of each service, our procedure will be to separately analyze the efficiency of the DMUs within each of two programs Joint advertising versus Service-Specific in pair-wise fashion. For the inefficient performers within each program, we determine the potential improvement in the number of recruitment contracts if they were operating efficiently, and then increase the inefficient units' output to this value. Thus, to make each DMU within a program efficient we remove the managerial or "technical" inefficiencies. We then combine the DMUs from the two programs and conducted another DEA analysis in which we analyzes the effects of advertising on the basis that all operations were conducted efficiently in each Service-Specific versus Joint comparison by ensuring that all operations (including advertising) were conducted efficiently. Essentially, this analysis puts the two programs (joint versus service specific advertising) on an equal footing of equal efficiency in managerial implementation.

The procedure we use is based on identifying a separate efficiency frontier for each service specific as well as a separate efficiency frontier for Joint advertising. All points within each service (as well as for Joint) are then projected onto their respective frontiers. The thus projected points can then be brought into play in another DEA analysis in the form of a pair-wise comparison with the similarly projected points for Joint advertising. In this manner, we can evaluate these collectives against a new frontier in order to determine the potentials of these Service-Specific and Joint in this pair-wise manner. In this manner, the "program" possibilities could thereby be evaluated free of

any shortcomings due to "managerial" inefficiencies. Then applying Mann-Whitney rank order statistics we can determine whether the Service-specific generally outranked the Joint advertising performances or vice-versa, in a statistically significant manner. Proceeding in this manner, we shall show that again Service-specific advertising was more efficient than Joint advertising.

Using the above-described DEA methodology, we can move on to evaluate the relative efficiencies of Service-specific versus Joint advertising for use in military recruiting. To do this, we will regard these as two separate programs and attempt to evaluate the relative efficiencies of these two different programs. A question that arises in such evaluations is how to distinguish between the potential of each program and the efficiency with which the two programs (Service-specific and Joint) are managed. To evaluate the comparative efficiencies of each advertising program, we will want to remove other inefficiencies that may cloud the differences (i.e., the relative efficiencies) of these two programs and consequently we shall distinguish between "managerial" and "program" efficiencies. Motivation for this distinction may be provided by noting that a good program might be badly managed and therefore appear to be less good than an inferior program that is well managed. One who is using averages or other measures of central tendency is especially vulnerable to making this error. It is an unfortunate reality that one must necessarily deal with data which intermingle these two types of inefficiencies, and one must construct a technique which remove the managerial efficiency differences prior to making judgments concerning the efficiency of the two programs.

5.3. *Eliminating "Managerial" Inefficiency to Focus on "Program" Efficiency*

The geometric figures that follow will help to portray how the comparative evaluations we are seeking are to be effected. Thus, consider Figure 2 as a starting point. Here we have portrayed the set of observations symbolized as "+" in this Figure. These observations are used to derive an efficient frontier in the manner indicated for DEA.

To start, assume we have applied a DEA model to the data of interest. Having thus effected our evaluations we can then construct the efficiency frontier and then project the inefficient units to the efficient frontier to obtain new points $(\hat{x}, \hat{y})$, which represent the coordinates of a point on the efficiency frontier corresponding to the projection of (x, y). To simplify, in our figure we deal with a single input x and a single output y. We then project the observed (x, y) values for the inefficient units represented by + onto the efficiency frontier in the manner indicated by the arrows. The new coordinates, $(\hat{x}, \hat{y})$, will have $\hat{x} = x$ in this single input case because we are adjusting toward efficiency by adjusting the output downward until it corresponds to the originally given input value.

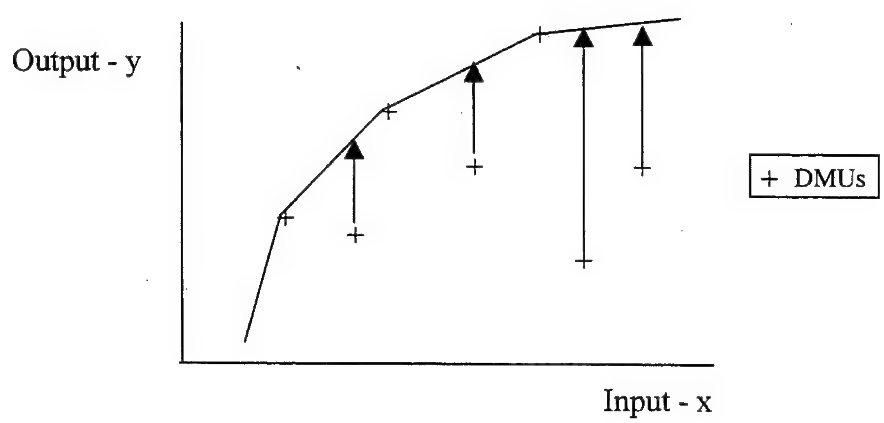


Figure 2: Adjustments to the Efficiency Frontier for Output

Applying this procedure separately to each of the two programs we are considering makes it possible to determine if (and by how much) one program is more efficient than the other. Figure 3 portrays the kind of situation we are considering in a very simple manner. Here we have two programs, which we label Program 1 and Program 2. The observed values of the former are represented by "o" and the observed values of the latter are represented by "+".

As can be seen, the observed behaviors represent a mixture of efficient and inefficient behavior in both programs with some of the "+" lying above some of the "o" and vice versa. Nevertheless, the efficiency frontier for Program 2 strictly dominates the efficiency frontier for Program 1. Hence it is possible to conclude that Program 2 is more efficient than Program 1 (One might miss this fact if one used averages or regressions).

The deviations from each of the two frontiers are said to represent "managerial inefficiencies" as they are not due to the Program itself. Unlike using an index number (or other) ratio averages, or unlike regressing output against input as in ordinary least squares, we can directly address the question of which "Program" is more efficient by eliminating managerial inefficiencies. We therefore now adopt the procedure initiated in Charnes, Cooper and Rhodes (1981) in order to obtain projections to determine "program efficiencies" – as distinct from the inefficiencies associated with the way these programs are managed.

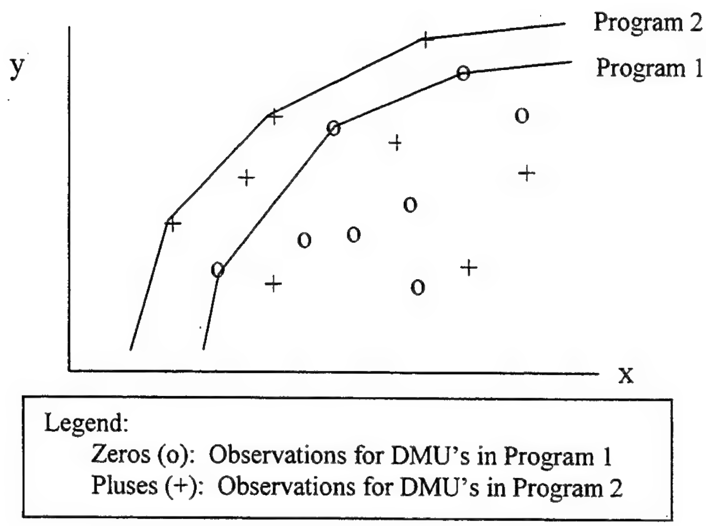


Figure 3: Program Efficiency

The procedure we will follow is similar to that outlined in Brockett and Golany (1996). First, each Program is evaluated separately by applying DEA to each sub-data set without reference to the other sub-data set for the other Program. Second, all points in each sub-data set are projected onto their respective efficiency frontier, thus eliminating managerial inefficiencies for that Program. Third, the resulting efficient points from both data sets are joined together and a new efficiency frontier is derived by applying DEA to the total data set consisting of these efficiency-adjusted values.

When this is all done, a new efficiency frontier is derived. If the two sub-data sets share the same efficiency frontier, then after projection, both should also lie on this single combined efficiency frontier and deviations from this combined efficiency frontier should be equally likely for each Program. On the other hand, if the individual efficiency frontiers are nested as in Figure 3, then the deviations from the combined efficiency frontier will be greater for the DMUs from the lower curve in Figure 3. The Mann Whitney test can again be used assess the statistical significance of the efficiency of one Program over another, as described in Brockett and Golany (1996).

Thus, in order to evaluate the hypothesis that the programs are equally efficient versus the alternative hypothesis that Program 1 is more efficient after removing (or controlling for) managerial inefficiencies, we do the following,

1. Run separate DEAs for each Program separately and determine the efficiency frontier for each.
2. Project each DMU to its Program Efficient Frontier, and combine all the thus projected DMUs into a single combined data set.
3. Run DEA again on the combined projected data set.
4. Rank the DMUs by their relative efficiency rating as obtained from the combined efficiency analysis of the efficiency-adjusted DMUs. All ties in efficiency ratings are assigned the mid-rank value.
5. Compute the sum of all rankings of one of the programs and compute the Mann-Whitney rank statistic U and its normal approximation Z_{test} as previously described.
6. We accept the null hypothesis that the programs are equally efficient if $-Z_{\alpha/2} \leq Z_{test} \leq Z_{\alpha/2}$.

When we apply this procedure to the Advertising Mix data to determine whether Joint or Service-specific advertising is the more efficient program for use in recruiting, we obtain the results given in Table 5.

	n_1	n_2	R	U	Z
Army	2148	384	2,386,445	746,413	**25.312**
Navy	2148	384	2,378,782	754,076.5	**25.893**
Air Force	2148	384	2,484,711	648,147	**17.865**
Marine	2148	384	2,453,064	679,794.5	**20.263**

Table 5: Hypothesis Testing Results Table using the Mann-Whitney Rank Statistic to Determine the Efficiency of Joint versus Service-Specific advertising Programs after Removal of Managerial Inefficiencies

As can be seen, the Z-values obtained in each study are extremely high indicating a high degree of confidence in our findings. Even the lowest Z-value, found in the Air Force at 17.865, reflects a very high level of statistical significance. Hence, for every service, we can reject the null hypothesis of equality of the two Programs, and conclude that from each services' viewpoint, a Service-specific advertising program is more efficient than a Joint advertising program once managerial inefficiencies are addressed. We conclude that any perceived dominance of Joint over service Specific is due to efficiency differences and not due to effectiveness of the advertising program.

6. Conclusions

It is to be noted now that our ability to distinguish program efficiencies by eliminating managerial inefficiencies, have led to results that differ markedly from the results reported by the Wharton Center for Applied Research (WCAR) and others. Whereas their results were inconclusive, our results yield high values of statistical significance favoring Service-specific over Joint advertising in every case. We have shown that due to managerial inefficiencies differences between the two advertising programs in the design of the Advertising Mix Test, one might be erroneously led to conclude that Joint is better that Service Specific, but this conclusion is reversed, both in

an analysis using only efficient units, and in an analysis wherein the inefficiencies have been removed. Service Specific advertising is statistically significant in it superiority over Joint service advertising.

References

Banker, R.D., A. Charnes and W.W. Cooper, "Some Models for Estimating Technical and Scale Inefficiency in Data Envelopment Analysis," *Management Science*, 30, 9 (1984), 1078-1092.

Brockett, P., and Golany, B., "Using Rank Statistics for Determining Programmatic Efficiency Differences in Data Envelopment Analysis", *Management Science*, Vol. 42, No. 3, March 1996, pp. 466-472.

Carroll, Vincent P., *DOD Advertising Mix Test: Comparison of Joint-Service with Service Specific Strategies and Levels of Funding,* Office of the Assistant Secretary of Defense Report, July 1987.

Charnes, A., W. W. Cooper, B. Golany, P. Brockett, *Critique of FY 1984 Advertising Mix Test of Wharton Center for Applied Research*, University of Texas at Austin, CCS Report 546, September 10, 1986

Charnes, A., W.W. Cooper and E. Rhodes, "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through," *Management Science*, 27, 6 (1981), 668-697.

Charnes, A., Cooper, W. W., and Rhodes, E., "Evaluating Program and Managerial Efficiencies: An Application of Data Envelopment Analysis to Program Follow Through", *Management Science*, 27, 1981pp. 668-697.

Cooper, W. W., L. M. Seiford, and K. Tone, *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, and References and DEA Solver Software*, Norwell, MA: Kluwer Academic Publishers, 1999

Dale, Charles, and Gilroy "The Changing Structure of the U. S. Economy: Its Effects on Army Enlistments," in Cutis Gilroy (ed.), *Army Manpower Economics*, Boulder, Colorado, Westview Press, 1986.

Daula, Thomas V., and D. Alton Smith, "Recruiting Goals and Enlistments in the U. S. Army," in Cutis Gilroy (ed.), *Army Manpower Economics*, Boulder, Colorado, Westview Press, 1986.

Dertouzos, James N., *Recruiter Incentives and Enlistment Supply*, RAND Note N-3065-MIL, The RAND Corporation, 1985.

Dertouzos, James N., *The Effects of Military Advertising: Evidence from the Advertising Mix Test*, RAND Note N-2907-FMP, The RAND Corporation, 1989.

Dertouzos, James, N., J. Michael Polich, *Recruiting Effects of Army Advertising*, RAND Report, R-3577-FMP, The RAND Corporation, 1989

Dertouzos, James, N., J. Michael Polich, *The Enlistment Bonus Experiment*, RAND Report, R-3353-FMP, The RAND Corporation, 1985

Frontier Analyst Professional, Banxia Holdings Limited, 1995-98.

Ganley, J. A. and J. S. Cubbin, *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis,* North-Holland, 1992.

Goldberg, L., "Recruiters, Advertising and Navy Enlistments," *Naval Research and Logistics Quarterly*, Vol. 29, No. 2, June 1982.

Goldberg, L. and Greenston, P., "Resource Analysis Study: How Educational Benefit Incentives and Alternative Levels and Resources Would Effect Recruiting Efforts Under Partial Mobilization", Economic Research Laboratory, Reston, VA, July 1986.

Kwinn, Michael J. Jr., "Evaluating Military Recruitment to Determine the Relative Efficiencies of Joint versus Service-Specific Advertising," Dissertation, University of Texas at Austin, 2000. Also available from University Microfiche, Inc., Ann Arbor, Michigan.

Norman, M and B. Stoker, *Data Envelopment Analysis" The Assessment of Performance,* New York, John Wiley & Sons, 1991.

Sohn, Sy, "Random Effects Meta Analysis of Military Recruiting", *Omega, International Journal of Management Science*, Vol. 24, No. 2, 1996, pp. 141-151.

Thomas, David A., "Data Envelopment Analysis Methods in the Management of Personnel Recruitment Under Competition in the Context of U. S. Army Recruiting," Dissertation, University of Texas at Austin, 1990. Also available from University Microfiche, Inc., Ann Arbor, Michigan.

US Government Accounting Office, *Advertising for Military Recruiting: How Effective Is It?,* FPCD-76-168, 1976.

US Government Accounting Office, *Improving the Effectiveness of Recruiting,* FPCD-15-69, 1976.

Warner, John T., "Military Recruiting Programs During the 1980s: Their Success and Policy Issues", *Contemporary Policy Issues*, Vol VIII, October 1990, pp. 47-67.